CLAIMS

1. A method for converting a voice signal (60) delivered by a source speaker into a converted voice signal (70) having acoustic features resembling those of a target speaker, comprising:

- the determination (1) of at least one function for transforming acoustic features of the source speaker into acoustic features similar to those of the target speaker, using voice samples from the source and target speakers; and

- the transformation (2) of acoustic features of the source speaker voice signal to be converted by applying the at least one transformation function,

characterized in that the transformation (2) comprises a step (44) for applying only a determined part of at least one transformation function to the signal to be converted.

2. A method according to claim 1, characterized in that the determination (1) of at least one transformation function comprises a step (20) for determining a model representing in a weighted manner common acoustic features of voice samples from the target speaker and from the source speaker on a finite set of model components, and in that the transformation (2) comprises:

- a step (36) for analyzing the voice signal to be converted, which voice signal being grouped into frames, in order to obtain, for each frame of samples, information relating to the acoustic features;

- a step (38) for determining an index of correspondence between the frames to be converted and each component of the model; and

- a step (40) for selecting a determined part of the components of the model according to the correspondence indices,

the step (44) for applying only a determined part of at least one transformation function comprising the application to the frames to be converted of the sole part of the at least one transformation function corresponding to the selected components of the model.

3. A method according to claim 2, characterized in that it additionally comprises a step (42) for normalizing each of the correspondence indices of the selected components with respect to the sum of all the correspondence indices of the selected components.

4. A method according to either claim 2 or claim 3, characterized in that it additionally comprises a step (43) for storing the correspondence indices

and the determined part of the model components, performed before the transformation step (44), which is delayed in time.

5. A method according to any one of claims 2 to 4, characterized in that the determination (1) of the at least one transformation function comprises:

- a step (4X, 4Y) for analyzing voice samples from the source and target speakers, grouped into frames in order to obtain acoustic features for each frame of samples from a speaker;

- a step (18) for the time alignment of the acoustic features of the source speaker with the acoustic features of the target speaker, this step (18) being performed before the step (20) for determining a model.

6. A method according to any one of claims 2 to 4, characterized in that the step (20) for determining a model corresponds to the determination of a Gaussian probability density mixture model.

7. A method according to claim 6, characterized in that the step (20) for determining a model comprises:

- a sub-step (22) for determining a model corresponding to a Gaussian probability density mixture, and

- a sub-step (24) for estimating parameters of the Gaussian probability density mixture from the estimation of the maximum likelihood between the acoustic features of the samples from the source and target speakers and the model.

8. A method according to any one of claims 1 to 7, characterized in that the determination (1) of at least one transformation function is performed based on an estimator of the realization of the acoustic features of the target speaker given the acoustic features of the source speaker.

9. A method according to claim 8, characterized in that the estimator is formed by the conditional expectation of the realization of the acoustic features of the target speaker given the realization of the acoustic features of the source speaker.

10. A method according to any one of claims 1 to 9, characterized in that it additionally includes a synthesis step (48) for forming a converted voice signal from the transformed acoustic information.

11. A system for converting a voice signal (60) delivered by a source speaker into a converted voice signal (70) having acoustic features resembling those of a target speaker, comprising:

- means (56) for determining at least one function for transforming acoustic features of the source speaker into acoustic features similar to those of the target speaker, using voice samples from the source and target speakers; and

- means (66) for transforming acoustic features of the source speaker voice signal to be converted (60) by applying the at least one transformation function,

characterized in that the transformation means (66) are adapted for the application only of a determined part of at least one transformation function to the signal to be converted (60).

12. A system according to claim 11, characterized in that the determination means (54) are adapted for the determination of at least one transformation function using a model representing in a weighted manner common acoustic features of voice samples from the source and target speakers on a finite set of components, and in that it includes:

- means (62) for analyzing the signal to be converted (60), which signal being grouped into frames, in order to obtain, for each frame of samples, information relating to the acoustic features;

- means (64) for determining an index of correspondence between the frames to be converted and each component of the model; and

- means (65) for selecting a determined part of the components of the model according to the correspondence indices,

the application means (66) being adapted for applying only a determined part of the at least one transformation function corresponding to the selected components of the model.